PARENTAL NETWORKS, SOCIAL CLOSURE, AND MATHEMATICS LEARNING: A TEST OF COLEMAN'S SOCIAL CAPITAL EXPLANATION OF SCHOOL EFFECTS*

Stephen L. Morgan
Harvard University

Aage B. Sørensen
Harvard University

Through an analysis of gains in mathematics achievement between the tenth and twelfth grades for respondents to the National Education Longitudinal Study of 1988, we examine Coleman's explanation for why Catholic schools apparently produce more learning than public schools. According to Coleman, Catholic schools benefit from larger endowments of social capital, generated in part through greater intergenerational social closure (i.e., dense network connections between the parents of students). Instead, we find that for public schools, social closure among parents is negatively associated with achievement gains in mathematics, net of friendship density among students. This evidence of a negative effect of parental social closure within the public school sector lends support to our alternative hypothesis that horizon-expanding schools foster more learning than do norm-enforcing schools. Moreover, this result renders social closure incapable of explaining any portion of the Catholic school effect on learning, even though within the Catholic school sector there is some evidence that social closure is positively associated with learning.

Coleman (1990) argues that the concept of social capital is valuable because it focuses analytic attention on the resources that inhere in social relationships, "those aspects of social structure . . . that can be used by the actors to realize their interests" (p. 305). Both the strength and weakness of this conceptualization rest in its ubiquitous nature—almost any aspect of social structure can be defined as social capital given an appropriate context for action.[1]

Further research on social capital must move in two directions. First, Coleman's theory of social capital must be confronted and reconciled with the prior theoretical work of other scholars. Portes (1998), Sandefur and Laumann (1998), and Woolcock (1998) have made progress in this regard. Second, empirical research must be mounted to judge the concept's potential for offering novel and parsimonious explanations for ob-

* Direct all correspondence to Stephen L. Morgan, Department of Sociology, Harvard University, William James Hall 5th Floor, Cambridge MA 02138 (smorgan@wjh.harvard.edu), or Aage B. Sørensen, Department of Sociology, Harvard University, William James Hall, Cambridge MA 02138 (abs@wjh.harvard.edu). An early version of this paper was presented at the meetings of the American Sociological Association in August, 1997. We are grateful for comments from Susan Dumais, Peter Marsden, William Morgan, Aaron Pallas, and Christopher Winship, and we acknowledge the financial support of the National Science Foundation through Morgan's graduate research fellowship. This research was also supported by a grant to Sørensen from the American Educational Research Association, which receives funds for its AERA Grants Program from the National Science Foundation and the National Center for Education Statistics (U.S. Department of Education) under NSF Grant #RED-9452861. Opinions reflect those of the authors and do not necessarily reflect those of the granting agencies.

[1] For example, see the contrasting context-specific definitions of social capital in Putnam (1993:167), Portes and Sensenbrenner (1993: 1323), and Burt (1997a:340). Because the concept is so broad, skeptics (e.g., Baron and Hannan 1994) view social capital as little more than a useful metaphor.

served behavior. Here, too, there has been recent progress. In research on outcomes for at-risk youth (Furstenberg and Hughes 1995; McLanahan and Sandefur 1994), the economic sociology of immigration (Portes 1995; Portes and Rumbaut 1996), and job promotion (Burt 1992, 1997a, 1997b, 1998; Podolny and Baron 1997), Coleman's theory has received direct attention in applications with real data. Nonetheless, Coleman's foundational example of a social capital effect—the effect of schools on student learning—has received comparatively little attention in empirical research. In this article, we demonstrate the need for and reward from evaluating Coleman's social capital theory of school effects.

COLEMAN AND THE CATHOLIC SCHOOL EFFECT ON LEARNING

Although it is customary to cite Coleman (1988a) as the urtext of his theory of social capital, Coleman first invoked the concept to explain differences in student learning across types of schools. Coleman and his colleagues became convinced that students who attend Catholic high schools learn more than similar students who attend public high schools (Coleman and Hoffer 1987; Coleman, Hoffer, and Kilgore 1982). Because Catholic schools spend less money per pupil, Coleman developed his theory of social capital to account for the existence of nonmonetary resources that give students in Catholic schools a learning advantage (also see Coleman 1987, 1988b, 1995).[2]

Coleman never developed an explicit mechanism to explain his core empirical finding that students enrolled in Catholic schools perform better on standardized tests of achievement than public school students with similar observed characteristics. In fact, the link by which endowments of social capital improve performance on standardized tests remained unspecified in all of his school-effects research. Nevertheless, he did delineate the two types of social capital that he believed combined to give Catholic school students a learning advantage: the ideology of the Catholic church and intergenerational social closure.

For Coleman, the sacred commitment of the Catholic church is a source of social capital for students enrolled in Catholic schools. He states, "the precept derived from religious doctrine that every individual is important in the eyes of God" leads educators to encourage all students to learn (Coleman 1990: 321). The returns that students draw on this form of social capital are the result of learning in response to achievement norms and teaching practices buttressed by religious conviction. This form of social capital cannot be created within a public school system, and Coleman recognized this impossibility. However, the second form of social capital from which Catholic school students purportedly benefit—social capital generated by social closure among parents in the school community—can be manufactured in the communities within which public schools are situated.

We evaluate whether the second of these two sources of social capital accounts for any portion of the Catholic school effect on achievement. Using data from the National Education Longitudinal Study (NELS) of 1988, we examine achievement gains in mathematics between the tenth and twelfth grades in an attempt to answer two related questions: Is social capital, in the form of social closure, associated with increased learning in mathematics? Can social closure explain a substantial portion of the Catholic school effect on learning?

To motivate our empirical analysis, we present hypotheses for positive and negative effects of social closure in the context of schools and communities. We introduce two different exemplars of school organization—the *norm-enforcing school* and the *horizon-expanding school.* Rather than rely on the capacity of closed social networks to enforce norms of diligence, horizon-expanding schools exploit a different type of social

[2] More recent analyses of the Catholic effect on achievement have obtained arguably more precise estimates using instrumental variable techniques (Figlio and Stone 1997; Hoxby 1996; Neal 1997), modeled heterogeneity of learning determinants with multilevel analysis techniques (Bryk and Raudenbush 1992 and citations therein), and elaborated the organizational features of effective schools in general (Chubb and Moe 1990; Lee, Smith, and Croninger 1997). Coleman's social capital explanation for the Catholic school effect on achievement has received no rigorous evaluation with the available survey data.

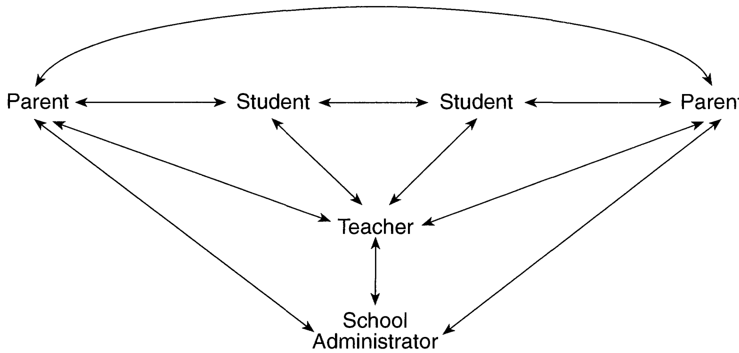

Figure 1. Social Connections in a Norm-Enforcing School

capital—information about opportunities in the extended social networks of parents and other adults.

NETWORK CONFIGURATIONS OF TWO MODELS OF SCHOOL COMMUNITIES

Coleman (1987, 1990) presents simple sociograms that vary in the connectedness of parents in order to show how social networks generate different amounts of social capital across school communities. He argues that when the parents of a group of students all know each other, valuable social capital resources accumulate in the ties among them that can promote student learning. The social network we present in Figure 1 is generally consistent with Coleman's idea of a closed functional community that facilitates learning. We call a school embedded in this pattern of network relations a norm-enforcing school.

The distinguishing feature of the social organization of a norm-enforcing school is the set of relationships forged among parents.[3] To effectively monitor the out-of-school behavior of their children and to exchange information, parents establish ties with the parents of their children's school friends. According to Coleman (1987, 1995), however, social closure among all adults in the school community can help maintain the value consistency of a functional community. Thus, teachers also form close ties with each other and with school administrators, cultivating communal organizational practices that foster learning (Bryk, Lee, and Holland 1993). And parents may also establish relationships with school administrators through involvement in parent-school organizations and volunteer work in order to influence school policy and monitor the performance of teachers.

Catholic schools are especially effective norm-enforcing schools because they can appropriate as social capital all of the social bonds maintained in a more encompassing functional community, the church. Above and beyond the beneficial effects of the church ideology, Coleman argues that the network connectedness fostered by communal religious observance creates an additional stock of social capital for Catholic schools.

Nonetheless, norm-enforcing schools have a dark side, similar to the "downside" of social capital noted by Portes and Landolt (1996). In their drive to maintain value consistency, norm-enforcing schools can become suffocating communities in which excessive monitoring represses creativity and exceptional achievement. The costs of social closure emerge in two forms: loss of autonomy and redundant information.[4] In theory,

[3] Nonetheless, as with every school the core of a norm-enforcing school is the network of relationships among students, teachers, and parents. Schools function best when students build strong bonds with their classmates, when teachers cultivate nurturing relationships with their students, and when parents establish close ties with teachers.

[4] For a discussion of loss of autonomy, see the "constraints on freedom" and "leveling pressures" subsections in Portes and Sensenbrenner (1993). For a discussion of opportunity costs

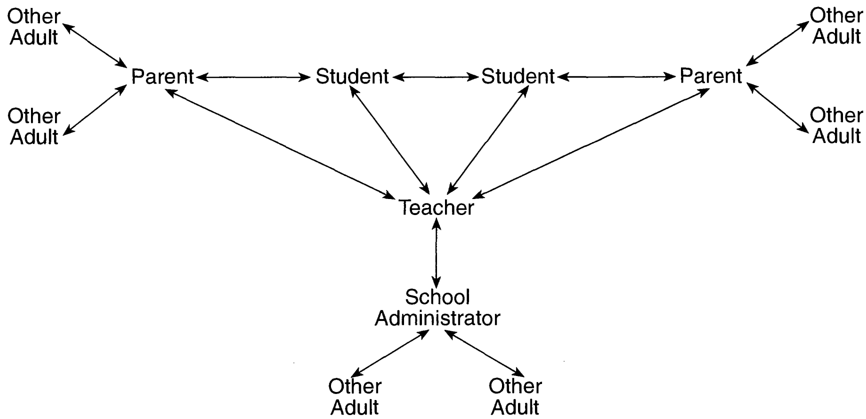

Figure 2. Social Connections in a Horizon-Expanding School

Coleman (1987) recognized these potential costs of social closure, noting that students who are products of such schools

> . . . may be unequipped to enter the heterogeneity and disorder of the larger society and thus either confined to the narrow frame within which they grow up or forced to suffer a serious culture shock when they leave their protected habitat. (P. 191)

Fortunately, the social control and information costs of social closure need not be highly correlated, as we show in the model of a horizon-expanding school presented in Figure 2. Like the organization of a norm-enforcing school, a horizon-expanding school is characterized by close ties among fellow students and their teachers, among fellow teachers, and among parents and teachers. But parents who send their children to horizon-expanding schools do not devote as much time to the cultivation of bonds with the parents of their children's school friends. Nor do they spend as much time developing bonds with school administrators. Through choice, often residential, parents select schools for their children where they expect school administrators to monitor teacher performance according to their wishes and other parents to reinforce achievement norms as they themselves would. Parents then spend relatively more time than parents of students in norm-enforcing schools investing in social capital outside of the immediate school environment.

We estimate random-effects analysis-of-covariance models (Bryk and Raudenbush 1992; Goldstein 1995; Longford 1993) to assess the relationship between peer and parental network configurations and students' achievement gains in mathematics between the tenth and twelfth grades. We also test Coleman's claim that Catholic school attendance and parental network effects on mathematics achievement are both positive and to some extent equivalent.

METHODOLOGY

Data and Variables

Data were drawn from the 1988 (base year) through 1992 (second follow-up) surveys of the National Education Longitudinal Study (NELS) of 1988 (U.S. Department of Education 1996). To reduce the dimensionality of the research questions to a manageable level and to fully exploit the panel nature of the data, the analytic sample chosen can be generalized only to tenth grade students enrolled in public and Catholic schools in 1990 *who also were enrolled in the eighth grade in any type of school in the United States in 1988.* As a result, the analytic sample cannot be generalized to the population that includes: (1) students who were enrolled in the tenth grade in 1990 in other types of schools (e.g., other religious and nonreligious private schools and home-schooled students) and (2)

from lost information, see the social network research on job search, competition, and promotion patterns (Burt 1992, 1997a, 1997b, 1998; Granovetter 1973, 1974; Podolny and Baron 1997).

Table 1. Means and Standard Deviations of Student-Level Variables: National Education Longitudinal Study of 1988

Variable	Mean	S.D.
Math Test Scores [a]		
IRT estimated number right (8th grade)	37.627	11.615
IRT estimated number right (10th grade)	45.404	13.236
IRT estimated number right (12th grade)	49.721	13.851
Race-sex Indicator Variables		
Black male	.041	.198
Hispanic male	.043	.203
Asian male	.015	.123
White female	.397	.489
Black female	.048	.214
Hispanic female	.043	.202
Asian female	.015	.121
Native American males and females	.006	.081
Family Structure Indicator Variables		
Mother only	.122	.327
Father only	.015	.122
Mother and stepfather	.090	.286
Father and stepmother	.018	.134
Other family type	.025	.155
Family data missing	.005	.073
Within-School Socioeconomic Status		
Mother's education (in years)	–.004	1.894
Father's education (in years)	–.112	2.197
SEI score of mother's occupation in 1992 (GSS 1989 coding)	–.820	10.412
SEI score of father's occupation in 1992 (GSS 1989 coding)	–.915	9.406
Family income in 1992 (ln)	–.006	.803
Math Courses Completed by the End of 10th Grade [b]		
At least one full year of algebra I and no full year of geometry, algebra II, or trigonometry	.215	.410
At least one full year of geometry and no full year of either algebra II or trigonometry	.336	.472
At least one full year of algebra II or trigonometry	.280	.449
Math Courses Completed by the End of 12th Grade [c]		
Algebra I	1.007	.636
Geometry	.756	.486
Algebra II	.593	.573
Trigonometry	.175	.353
Precalculus	.149	.350
Calculus	.102	.307

(Table 1 continued on next page)

(Table 1 continued from previous page)

Variable	Mean	S.D.
Realized Sample Selection Probabilities [d]		
Same school for 10th and 12th grade and completed all math tests (SSMAT)	.671	.118
Same school for 10th and 12th grade but did not complete all math tests (SSNMAT)	.173	.054
Changed schools between 10th and 12th grade and completed all math tests (CSMAT)	.033	.017
Changed schools between 10th and 12th grade but did not complete all math tests (CSNMAT)	.043	.027
In school all NELS years, not in usual grade, but completed all math tests (NIGMAT)	.006	.010
In school all NELS years, not in usual grade, and did not complete all math tests (NIGNMAT)	.005	.007
Dropped out at least once but completed all math tests (DOMAT)	.019	.023
Dropped out at least once and did not complete all math tests (DONMAT)	.039	.049
Follow-up status unknown (STATUK)	.010	.011

Notes: N = 9,241 for all variables except math-course-taking. The three variables for math courses completed by the end of tenth grade have valid Ns of 9,056, 8,919, and 8,742. All of the twelfth-grade math-course-taking variables have valid Ns of 8,837. Data are weighted by a within-school student weight multiplied by a school weight.

[a] Scores were estimated using item response theory. See text below.

[b] Data came from student reports in first NELS follow-up, 1990.

[c] Data were taken from transcript files and are measured in Carnegie units.

[d] Probabilities were estimated from the base-year sample.

students who were enrolled in the tenth grade in 1990 but who were not enrolled in the eighth grade in 1988 in the United States (e.g., those who were held back in the tenth grade, those who skipped the ninth grade, and those who emigrated to the United States after 1988). The NELS data set does contain sample members that can be used to make generalizations to some of these groups of excluded tenth graders, but we chose not to include them in our analysis.

Our models of mathematics gains are estimated on the sample of 9,241 respondents meeting the above criteria, nested within 898 schools. Appendix A provides a detailed description of how the analytic sample was formed and how selection into the final analysis sample is modeled (i.e., selection into the complete data subsample).

Student-level variables used for the analysis of math achievement are presented in Table 1. Corresponding school-level variables are presented in Table 2. For all three waves of the NELS, raw scores on multiple test forms of mathematics were transformed by contractors to the National Center for Education Statistics to common scales using item response theory (Hambleton, Swaminathan, and Rogers 1991; Lord 1980; Rock and Pollack 1995). All other student-level variables are self-explanatory. White males and mother-father families are designated as respective reference categories for the indicator variables characterizing race-gender and family type.

The school-level variables require more explanation. To investigate the social capital hypotheses developed by Coleman, a short battery measuring social networks was included in the second follow-up parent questionnaire. Parents were asked to write down the names of their teenager's five closest friends and then indicate whether each friend attended the same school as their teenager. *Friends in school* was computed as the mean for each school of all parents' sums of positive responses to these prompts and thus ranges from 0 to 5. Parents were next asked to indicate whether they personally knew at least one parent of each of their teenager's

Table 2. Means and Standard Deviations of School-Level Variables: National Education Longitudinal Study of 1988

Variable	Mean	S.D.
Social Closure and Parental Involvement		
Social closure around school	3.067	.742
Number of a student's five closest friends that attend the same school	3.312	.734
Number of a student's five closest friends' parents known by a student's parents	3.121	.748
Parents work together supporting school policy	2.712	.336
Parents have adequate say in school policy	2.563	.356
Catholic school indicator variable	.059	.237
Urbanicity-Region (for Public Schools)		
Suburban, Northeast	.088	.284
Suburban, South	.123	.329
Suburban, West	.073	.260
Urban, Midwest	.036	.187
Urban, Northeast	.041	.199
Urban, South	.085	.279
Urban, West	.047	.212
Rural, Midwest	.102	.303
Rural, Northeast	.041	.199
Rural, South	.138	.345
Rural, West	.046	.209
Socioeconomic Status		
Mother's education (in years)	12.999	1.312
Father's education (in years)	13.479	1.685
SEI score of mother's occupation in 1992 (GSS 1989 coding)	45.286	7.102
SEI score of father's occupation in 1992 (GSS 1989 coding)	45.372	7.498
Family income in 1992 (ln)	10.450	.551

Notes: N = 898. For each variable, values are within-school weighted means of student characteristics. The mean and standard deviation of each variable across schools is weighted by a school-level weight (the within-school unweighted mean of student weights).

five closest friends. *Parents know parents* is the mean for each school of all parents' sums of responses to this second set of prompts and also ranges from 0 to 5. *Social closure around school* is the square root of the product of these two dimensions and again ranges from 0 to 5.

We utilize two other school-level variables. *Parents have adequate say* is the mean for each school of parental agreement (on a four-point scale) with the statement, "Parents have an adequate say in setting school policy." Similarly, *parents work together* is the mean for each school of parental agreement with the statement, "Parents work together in supporting school policy." Both response scales were initially coded 1 for "strongly agree" through 4 for "strongly disagree." For our analysis, the codes were reversed prior to calculating the school means. All other school-level variables are self-explanatory. There is no indicator variable for suburban Midwestern public schools because they are designated as the reference category for the urbanicity-region of public schools.

Methods and Models

Because the NELS is a multi-stage stratified random sample of students nested within schools, we estimate random-effects analysis-of-covariance models with both student-level and school-level weights and disturbance terms. Equation 1 is a general representation of the models that we estimate:

$$Math12_{ij} - Math10_{ij} = d + (\lambda - 1)Math10_{ij} + cC_j + \mathbf{a}'\mathbf{X}_j + \mathbf{b}'\mathbf{X}_{ij} + u_j + e_{ij}, \quad (1)$$

where $Math12_{ij}$ is the twelfth-grade math test score, $Math10_{ij}$ is the tenth-grade math test score; d is a constant that is an estimate of the mean gain in math achievement for public school students; c is an estimate of the Catholic school treatment effect, because C_j is an indicator variable for Catholic schools; $\mathbf{a}$ is a vector of school-level fixed effects on math achievement, because $\mathbf{X}_j$ is a vector of centered school-level characteristics that vary only over the j schools in the sample; $\mathbf{b}$ is a vector of student-level fixed effects on math achievement, because $\mathbf{X}_{ij}$ is a vector of centered student-level characteristics that vary over the i respondents in each j school; u_j is a mean zero school-level error term, and e_{ij} is a mean zero student-level error term. We estimated fully specified versions of equation 1 with iterative generalized least squares and a robust variance estimator, as implemented by the program MLwiN (Version 1.02). Appendix B discusses the limitations of the model represented by equation 1, including a critical assessment of alternative specifications.

RESULTS

Main Findings

Table 3 presents coefficients of primary interest from four models in the form of equation 1. Coleman's main hypothesis is that closed functional communities foster student learning. Model 1 is a first attempt to evaluate this hypothesis, as it predicts student gains in mathematics between the tenth and twelfth grade from three school-level variables that measure social closure in the communities that surround schools, cooperative parental involvement in schools, and parental satisfaction with input into school policy.

For Coleman's theory, the most relevant predictor is *social closure around school.* According to Coleman, social closure exists around a school when all of the students' close friends attend the school and all of the students' parents know each other, as is the case in the norm-enforcing model of a school presented in Figure 1.[5]

We treat the other two school-level predictors as covariates of secondary interest. Norm-enforcing schools with high social closure should have high values for *parents work together*, even though it is theoretically possible that parents in closed functional communities may spend so much time socializing outside of schools that they do not collectively shape school policy. Likewise, norm-enforcing schools should have high values for *parents have adequate say* because parents in communities surrounding norm-enforcing schools feel that they are an integral part of a school's norm-maintenance system. However, horizon-expanding schools and other types of schools may also have high values on this variable. If parents choose schools through residential choice that match their school-policy preferences, then parents may feel that they have an adequate say in setting school policy even though they have never attempted to directly influence school policy.

Model 1 specifies these three school-level characteristics as independent variables along with most of the other covariates presented in Tables 1 and 2. Math gains between the tenth and twelfth grades is the dependent variable, and the core learning model specifies tenth grade math score as an independent variable, as in equation 1. In all models, the tenth-grade math score has a negative relationship with gains in math achievement between the tenth and twelfth grade. The negative coefficient indicates that there is either

[5] As a source for direct measurements of social closure, the NELS data are better than other existing data that we know of but still are inadequate. The NELS data do not furnish information on how many close friends students have, only on the number of a student's five closest friends who attend the same school. The questions on parental networks do not elicit the number and location of bonds that parents maintain with other adults, only the number of a student's five closest friends' parents that are known by a student's parent. Specialized data sets with more detailed social network information exist (e.g., Stanton-Salazar and Dornbusch 1995), but they appear to be too small to yield any generalizable conclusions, especially across schools.

Table 3. Coefficients from the Regression of Math-Score Gains between the Tenth and Twelfth Grades on Variables Indicating Social Closure: National Education Longitudinal Study of 1988[6]

Independent Variables	Model 1	Model 2	Model 3	Model 4
FIXED EFFECTS				
Constant	4.401	4.307	4.317	4.307
School-Level Variables[a]				
Catholic school	—	1.624*** (.316)	1.494*** (.326)	1.731*** (.321)
Social closure around school	.046 (.125)	—	.077) (.125)	—
Parents work together supporting school policy	.538 (.356)	—	.195 (.366)	—
Parents have adequate say in school policy	.331 (.324)	—	.413 (.328)	.536* (.238)
Friends in school	—	—	—	.379** (.125)
Parents know parents	—	—	—	–.314* (.133)
Student-Level Variables[b]				
IRT math score in 10th grade	–.105*** (.006)	–.106*** (.006)	–.106*** (.006)	–.107*** (.006)
RANDOM EFFECTS				
School-level variance	.982 (.234)	.944 (.224)	.930 (.228)	.890 (.232)
Student-level variance	28.123 (.853)	28.115 (.853)	28.108 (.851)	28.112 (.852)
–2 log-likelihood	57,645	57,632	57,627	57,618

Notes: N = 9,241 students in 898 schools. Data are weighted at both the student and school levels. Robust standard errors (in parentheses) are calculated with MLwiN's implementation of White's sandwich variance estimator.

[a] Additional school-level covariates include urbanicity-region for public schools, mothers' years of education and fathers' years of education and their SEI scores, and logged family income (see Table 2). All school-level variables are entered as grand-mean centered fixed effects, except the Catholic school indicator variable, which is entered as an uncentered fixed effect.

[b] Additional student-level covariates include race-sex and family structure, mothers' years of education and fathers' years of education and their SEI scores, logged family income, and a polynomial coding (orthogonal and of degree three) of the probability of remaining in the same school for the tenth and twelfth grades and completing all math tests. All student-level variables are entered as grand-mean centered fixed effects with the exception of the socioeconomic status covariates which are entered as group-mean centered fixed effects.

$^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$ (two-tailed tests)

regression toward the mean between the tests as a result of measurement error or that math learning is governed inherently by a concave growth function. The negative coefficient, although highly significant in a statistical sense, is not large in comparison with those typically obtained in the estimation of similar learning-gains models with alternative

[6] We have added asterisks to our regression tables (Tables 3, 4, and 5) to accommodate *ASR*/ASA style guidelines. We did not include asterisks in our original manuscript because (1) we are not fans of frequentist tests of point-value null hypotheses, (2) some readers mistake asterisks for substantive importance, and (3) asterisks are redundant when standard errors are provided.

data (e.g., Coleman and Hoffer 1987), suggesting that ceiling effects and regression toward the mean are unlikely to seriously bias the results for an analysis of math gains using the NELS data.

The constant for Model 1 is 4.401. Because all independent variables are centered around their mean values, the constant indicates that the average student gained approximately 4.5 points of math ability (or skill) between the tenth and twelfth grades. Since the math scores were transformed to estimated number right scores with item response theory, the scale of math gains is arbitrary. As a result, 4.401 points cannot be easily interpreted. Strict comparisons with the standard deviation of math scores in the tenth grade (13.236) are unwise, because the tenth-grade standard deviation is inflated by measurement errors of various types. Nonetheless, we can safely note that students do not gain tremendous amounts of mathematics skill in the last two years of high school, as their average math gain is only about one-third of the standard deviation in IRT math scores for tenth-graders.

Model 1 does not support social closure explanations for mathematics learning—the estimated coefficient for *social closure around school* is nearly 0.[7] The coefficients for *parents work together* and *parents have adequate say* are positive but less than twice their standard errors. To some extent, the large standard errors for these two regression coefficients result from the collinearity of the two variables, as their correlation coefficient is .688.[8]

[7] The support of Carbonaro (1998) for a positive effect of social closure on mathematics achievement is based on coefficients from cross-sectional models alone. When he adds prior achievement to his models as an independent variable, support for the positive effect vanishes. We find no justification in this result for Carbonaro's (1998:305) claim that this vanishing effect is evidence that social closure "operates primarily through students' prior achievement."

[8] These two school-level predictors have not explained away a social closure association with math achievement gains. *Parents work together* and *parents have adequate say* have zero-order correlation coefficients with social closure of only –.001 and –.072, respectively.

Model 2 substitutes a Catholic school indicator variable for the social closure, parental cooperation, and parental satisfaction variables that were included in Model 1. The coefficient for the Catholic school effect is 1.624, suggesting that Catholic school students learn approximately 38 percent more mathematics between the tenth and twelfth grades than do public school students. Model 2 serves as the baseline estimate of the Catholic school effect.

Coleman's theory of social capital predicts that a substantial portion of the Catholic school effect on achievement can be attributed to the greater parental involvement and social closure of the communities that surround Catholic schools. Model 3 evaluates this prediction, adding the three school-level variables of Model 1 to Model 2.

Model 3 provides almost no support for the parental involvement and social closure portions of Coleman's explanation of the Catholic school effect—only eight percent of the baseline Catholic school effect is explained by the inclusion of the three additional covariates. Moreover, the point estimate of the coefficient for *social closure around school* is still close to zero. Model 3 does not support the social closure portion of Coleman's explanation for the Catholic school effect on learning or the broader hypothesis that the social closure that characterizes norm-enforcing schools increases student learning.

Model 4 tests whether the alternative hypothesis—that horizon-expanding high schools foster more learning than do norm-enforcing high schools—is supported by the NELS data. The measure of social closure included in Model 3 is inappropriate for this evaluation because it combines into one measure two separate dimensions of social closure that may operate in opposite directions. As Figure 2 indicates, students in horizon-expanding schools are closely tied to each other, but their parents are not tied to each other. Model 4 specifies each underlying dimension of social closure separately, as school means of *friends in school* and *parents know parents*. Model 4 also retains *parents have adequate say* as a predictor but discards the *parents work together* variable because of concerns about multicollinearity.

Model 4 provides some support for the hypothesis that horizon-expanding schools produce more learning. The two dimensions of social closure operate in opposite directions, thus explaining why the coefficients for *social closure around school* in Models 1 and 3 are so small. Schools in which students are closely tied produce more learning. And net of this effect, schools around which parents are closely tied produce less learning. In addition to these countervailing network effects, *parents have adequate say* continues to have a positive effect, indicating that schools with policies that satisfy parental preferences also produce more learning.[9]

All three net effects, while not overwhelming in size, do seem to be substantively meaningful. An increase of one standard deviation in parental density decreases student learning by 5.5 percent (–.314 [.748 / 4.307] = –.055). Likewise, simultaneous increases of one standard deviation in student friendship density and parental satisfaction with school policy increase student learning by 6.5 and 4.4 percent, respectively.

The point estimate of the Catholic school effect for Model 4 is slightly larger than the baseline estimate of Model 2, indicating that the other school-level predictors of math achievement do not, taken together, explain any substantial portion of the Catholic school effect on learning. If anything, Model 4 suggests that the Catholic school effect is slightly larger than is suggested by the baseline estimate of Model 2.

Possible Variation in Effects across School Sectors

Partly because of the ambiguity in interpretation for the slight decreases and increases in the Catholic school effect from Models 2 through 4, we estimated two additional models that allow the social closure and parental involvement effects to vary across school sectors. Model 5, presented in Table 4, adds three interaction terms between the Catholic school indicator variable and *social closure around school*, *parents work together*, and *parents have adequate say* to the variables

[9] The coefficient is larger and its standard error smaller in this model partly because *parents work together* is not included.

Table 4. Coefficients from the Regression of Math-Score Gains between the Tenth and Twelfth Grades on Variables Indicating Social Closure across School Sectors: National Education Longitudinal Study of 1988

Independent Variable	Model 5	Model 6
FIXED EFFECTS		
Constant	4.316	4.305
School-Level Variables[a]		
Catholic school	1.786*** (.416)	1.693*** (.388)
Social closure around school	.026 (.130)	—
Social closure around school × Catholic school	.705 (.375)	—
Parents work together	.259 (.383)	—
Parents work together × Catholic school	–.865 (1.163)	—
Parents have adequate say	.341 (.339)	.509* (.250)
Parents have adequate say × Catholic school	.847 (1.188)	.328 (.823)
Friends in school	—	.383** (.131)
Friends in school × Catholic school	—	.101 (.390)
Parents know parents	—	–.368* (.142)
Parents know parents × Catholic school	—	.584 (.339)
Student-Level Variables[a]		
IRT math score in 10th grade	–.106*** (.006)	–.107*** (.006)
RANDOM EFFECTS		
School-level variance	.917 (.226)	.878 (.230)
Student-level variance	28.109 (.851)	28.114 (.852)
–2 log-likelihood	57,624	57,615

Notes: N = 9,241 students in 898 schools. Data are weighted at both the student and school levels. Robust standard errors (in parentheses) are calculated with MLwiN's implementation of White's sandwich variance estimator.

[a] Additional school-level covariates and student-level covariates are the same as for Table 3.

*$p < .05$ **$p < .01$ ***$p < .001$ (two-tailed tests)

included in Model 3. Similarly, Model 6 adds three interaction terms between the Catholic school indicator variable and *parents have adequate say, friends in school,* and *parents know parents* to the variables included in Model 4.

There are too few Catholic schools in the NELS sample and too little variation in patterns among them to provide statistically reliable estimates of how social closure and parental involvement effects vary across school sectors. Nonetheless, Models 5 and 6 highlight some interesting patterns in the observed data and suggest that, at a minimum, the social-closure, parental-involvement, and parental-satisfaction effects estimated in Models 3 and 4 are dominated by patterns that exist among public schools.

In Model 5, the point estimate of the *social closure around school* main effect is nearly zero. However, the coefficient for the *social closure around school* by *Catholic school* interaction is .705. While this point estimate of the interaction effect should be regarded with some caution because it is slightly less than twice the size of its standard error, it suggests that social closure has a positive association with learning *within* the Catholic school sector even though it has no effect within the public school sector.

The coefficients for the interactions of *parents work together* and *parents have adequate say* with *Catholic school* indicate that Model 5 is not well specified. The zero-order correlation between *parents work together* and *parents have adequate say* is even higher among Catholic schools alone at .706. This high level of collinearity contributes to the large standard errors of the interaction terms and likely has produced the nonsensical coefficient estimates of opposite sign.[10]

Model 6 breaks social closure into its two dimensions and drops the *parents work together* variable. Model 6 provides some clarification of the suggestive social closure differences of Model 5. While the coefficient for the *friends in school* by *Catholic school* interaction is small, the analogous coefficient for the *parents know parents* by *Catholic school* interaction is large enough to be substantively meaningful and also large enough to approach conventional statistical significance. Model 6 suggests that friendship density fosters learning in both public and Catholic schools. However, parental density limits student learning in public schools while possibly increasing learning in Catholic schools. These findings suggest that the most effective public schools are characterized by horizon-expanding patterns of social relations while the most effective Catholic schools are characterized instead by alternative norm-enforcing patterns of social relations.

Course-Taking as a Possible Intervening Mechanism

Virtually all past research on the Catholic school effect by Coleman and his colleagues and by their critics has maintained that Catholic schools achieve much of their learning advantage by requiring all students to learn a more challenging curriculum. Yet, when Catholic school students are compared with public school students from the highest curriculum tracks, much of the Catholic school effect vanishes.

Should we therefore estimate models that "control" for math-course-taking? There is no simple answer to this question, as a clear answer can only be offered if the causal ordering of social closure, school-sector choices, curriculum-development decisions, and course-taking choices is known.[11] Without knowing for certain how parents choose schools for their children and how teachers and school administrators make curriculum and tracking decisions, models with math-course-taking as covariates do not necessarily clarify conclusions. Despite these reser-

[10] See Winship (1998) for an explanation of the problems that multicollinearity produces. There is another interesting fact revealed by zero-order correlations. Whereas neither *parents work together* nor *parents have adequate say* has a substantial correlation with *social closure* for public schools, both variables have small positive correlations, .209 and .162 respectively, with *social closure around school* as measured in Catholic schools. This difference in the associations between the independent variables across the two school sectors contributes to the messy appearance of the interaction effects.

[11] See Morgan (1983) for a relatively non-partisan discussion of the endogeneity of both school sector and curriculum track with respect to family background.

vations, we estimated the following models because they document the patterns that exist in the NELS data. We will resist the temptation to take a strong position on what they actually mean.

The NELS data include detailed course-taking information in a supplementary transcript file constructed after the 1992 follow-up was completed. Information from this file is incomplete, but it provides a record of course-taking by the end of high school. And in conjunction with self-reported course-taking from the tenth-grade questionnaire, adding all of the math-course-taking variables from Table 1 to our models gives an adequate "control" for differential math-course-taking between the tenth and twelfth grades.

Table 5 presents two different versions of our preferred Model 4 from Table 3. In the first column, we reestimate Model 4 in its original form using the subset of respondents for whom complete math-course-taking data are available. Of the 9,241 respondents for whom all other models are estimated, 8,322 (or 90 percent) have complete math-course-taking data. Rather than attempt to impute values for the missing data on all nine variables, we simply reestimated the original Model 4 on the 8,322 respondents to demonstrate that missing-data patterns appear to be largely random with respect to the independent variables and the specification of the model. Differences between the coefficients from both versions of Model 4 are smaller than the standard errors of the original coefficients.[12]

In the second column of Table 5, we report coefficients from an augmented Model 4 that includes math-course-taking covariates (and that should be directly compared only with the coefficients presented in the first column of Table 5). The additional covariates are associated with achievement gains just as one would expect (but, to save space, are not reported). Net of the tenth-grade course-taking covariates, each Carnegie unit (a standard year of material) of calculus, precalculus, trigonometry, algebra II, and algebra I is associated with achievement gains on the mathematics tests of 2.89, 2.98, 2.11, 2.15, 1.71, and .91 points, respectively. Taken together, the addition of the nine math-course-taking variables yields a likelihood-ratio test statistic of 1,032.12, a highly significant departure from the mean of a chi-squared distribution with nine degrees of freedom.

Consistent with past research on the Catholic school effect, the curriculum covariates

Table 5. Preferred Model 4 Estimated for the Subset of Respondents with Complete Data for Math Course-Taking

Independent Variable	Model 4 without Math-Course Covariates	Model 4 with Math-Course Covariates
FIXED EFFECTS		
Constant	4.332	4.362
School-Level Variables[a]		
Catholic school	1.673***	.645*
	(.333)	(.323)
Parents have adequate say	.607*	.384
	(.247)	(.233)
Friends in school	.462***	.414**
	(.134)	(.137)
Parents know parents	–.301*	–.334*
	(.146)	(.144)
Student-Level Variables[b]		
IRT math score in 10th grade	–.107***	–.223***
	(.007)	(.008)
RANDOM EFFECTS		
School-level variance	.913	1.094
	(.237)	(.237)
Student-level variance	27.631	24.189
	(.841)	(.765)
–2 log-likelihood	51,770	50,738

Notes: N = 8,322 students in 850 schools. Data are weighted at both the student and school levels. Robust standard errors (in parentheses) are calculated with MLwiN's implementation of White's sandwich variance estimator.

[a] Additional school-level covariates are the same as for Table 3.

[b] Additional student-level covariates are the same as for Table 3 along with nine math-course-taking variables entered as grand-mean centered fixed effects.

*$p < .05$ **$p < .01$ ***$p < .001$ (two-tailed tests)

[12] This comparison provides no evidence that the data are not missing as a function of the dependent variable. The small increase in the constant and the commensurate decrease in the Catholic school effect suggest that the data are missing, to some small degree, as an inverse function of the dependent variable. We do not believe that more complicated models are justified for this mostly exploratory portion of our analysis.

decrease the Catholic school association with achievement gains by more than 60 percent. However, the social closure associations with achievement gains are nearly unaltered. These patterns emerge because course-taking is strongly related to school sector but not to either dimension of social closure.

DISCUSSION

Summary of Main Findings

In public high schools, the density of student friendship networks increases mathematics learning while the density of parental networks decreases it. In combination, differences in social closure among public schools have no association with differences in learning. Therefore, social closure cannot explain away any substantial portion of the observed Catholic school effect on learning.

These findings suggest that the apparent superiority of Catholic schools and the role of social closure in promoting student learning are not as closely related as Coleman's empirical findings led him to believe. In contrast to his basic hypotheses, our findings lead us to conclude that the benefits offered by the typical network configurations of horizon-expanding schools outweigh those of norm-enforcing schools, at least in the public school sector.

Network Configurations and Learning Mechanisms

By what mechanism can student and parental social network configurations influence student learning? We assume that learning depends on ability, effort, and opportunity. Among these determinants, network properties are most likely to affect learning by increasing student effort and opportunities to learn.

To some extent, Coleman relied on these same mechanisms in his explanations, arguing that Catholic schools are embedded in communities with stronger achievement norms that compel student diligence and thereby increase *student effort*. In support of this position, he presented evidence that students who attend Catholic schools have higher educational expectations and have parents who expect higher achievement from their children. To extend his argument to all schools, Coleman claimed that school communities rich in the social capital generated by social closure can better enforce achievement norms to bolster student effort.

A limitation of Coleman's generalized argument is that closed functional communities do not always construct and maintain norms that direct student effort toward learning. For example, one motivation for child labor laws and the establishment of large school districts at the beginning of the twentieth century was to emancipate children from achievement norms directed toward workplace behavior that were too strongly maintained in small communities (Tyack 1974). More recently, research on immigrant communities that are high in social capital and concerned about the intergenerational continuity of their enclaves suggests that these communities may subvert students' efforts in school by demanding community devotion (Portes and Rumbaut 1996). Accepting Coleman's general claim that adolescents benefit from parental density in a community, Wilson (1996) cites ghetto behavior in the 1990s as a counterexample, claiming that "social integration may not be beneficial to adolescents who live in neighborhoods characterized by high levels of individual and family involvement in aberrant behavior" (p. 62).

Especially (but not exclusively) in these contrary situations, abundant information contacts with the society outside of the school community may increase student effort, as we suppose is the case for students enrolled in horizon-expanding schools. Heterogeneous flows of information into a community enable parents and other adults to increase student effort by directing students' attention toward higher standards of achievement, successful role models, and desirable positions in society.

The second basic mechanism through which networks may affect student learning is *exposure to opportunity*. Within schools, opportunities to learn are a function of the instruction offered to students. Norm-enforcing schools are more responsive to parents' curriculum desires, and parental control may increase opportunities for student learning. But parents do not always know best. Indeed, in *The Adolescent Society*, Coleman (1961) wrote:

> Parents are often obsolescent in their skills, trained for jobs that are passing out of existence, and thus unable to transmit directly their accumulated knowledge. They come to be 'out of touch with the times,' and unable to understand, much less inculcate the standards of a social order that has changed since they were young. (P. 2)

For the same reasons that some communities rich in social capital may maintain achievement norms that are not directed toward learning, the parents associated with norm-enforcing schools may favor traditional or basic curricula that limit student potential. By contrast, we expect that parents who send their children to horizon-expanding schools expect curricular decisions to be driven by school administrators with substantial professional expertise.

Even beyond classroom instruction, there are opportunities for learning outside of schools. Students may acquire some basic knowledge and develop learning strategies by participating in adult-led organizations (e.g., church groups, scout troops, musical ensembles, and athletic teams) and adult-sponsored informal activities (e.g., hiking trips and visits to museums). The adult leaders or sponsors of such organizations and activities may be a student's own parents. But for most students, they are their friends' parents and their parents' friends.

Our findings do not allow us to identify whether the network effects on learning that we observe operate primarily through any subset of these mechanisms. Nonetheless, the finding that horizon-expanding public schools produce more learning can be interpreted as support for at least one of the two following assertions: (1) Exposure to the wider society within which local school communities are embedded increases students' efforts to learn; (2) social closure among parents limits access to informal learning opportunities provided by information flows from the wider society.

This interpretation of our findings as evidence primarily for learning effects that are responses to increased effort and opportunity is consistent with social network research on job promotion, competition, and entrepreneurship. With respect to information acquisition, close friends and friends' parents are structurally redundant. Most privileged information from friends' parents can be routed directly through a student's friends. However, access to information outside a student's peer network is enhanced through the ties that a student's parents build to information sources that are independent of a student's peer network. In social network terminology, students benefit more from the maintenance of weak ties to their parents' friends through their parents than from the deepening of ties they already have with their friends' parents through their friends.[13]

Social Capital and the Catholic School Effect on Learning

The most powerful prima facie explanation of the Catholic school effect on learning is that Catholic schools force all students through more challenging curricula by offering only college preparatory courses. Our findings support this explanation, as our mathematics course-taking model demonstrates that 60 percent of the baseline Catholic school effect can be accounted for by covariates that measure differential course-taking patterns. We believe, however, that course-taking is endogenous with respect to other basic processes, including adherence to norms buttressed by the ideology of the Catholic church. The more ambitious curricula of Catholic schools is not a sufficient explanation for the Catholic school effect on learning.

The value of the concept of social capital is that it can be used to map properties of social structure to sets of mechanisms that generate action. These mechanisms can then be presented in a common theoretical framework and jointly evaluated in an empirical analysis. Coleman and his colleagues insisted that students who attend Catholic schools learn more because they benefit from larger endowments of social capital. Unfortunately, they did not have appropriate data to test the mechanisms of Coleman's social capital explanation of school effects.

---

[13] In other words, students benefit from structural holes in the networks of parents that surround their schools (Burt 1992). Students also benefit from sufficiently strong and supportive relationships with their schoolmates (Burt 1997a, 1997b; Podolny and Baron 1997).

Now that appropriate data are available, a viable empirical analysis can be mounted. Based on our findings, we conclude that net associations between learning and the network components of social closure suggest that *if* the Catholic school effect on achievement is the result of a greater relative endowment of social capital, this capital must be the appropriable norms of the Catholic church.[14]

These norms may strengthen the common curriculum offerings, foster strict but cooperative teaching practices, and increase students' efforts. While the maintenance and enforcement of these norms may be a function of the closure of the networks of parents that surround Catholic schools, extrapolation to the public sector of any positive effects of parental social closure is unwarranted. The public sector lacks an other-worldly institution from which to appropriate indisputable norms in order to overcome the inherent costs of parental social closure alone.

[14] The main competing explanation of the Catholic school effect is that it is an artifact of aggregated patterns of individual self-selection (see Murnane, Newstead, and Olsen 1985). More research on this possible explanation is needed, but we do not believe that the NELS data alone are up to the task. For recent attempts that use instrumental variables, see Hoxby (1996) and Figlio and Stone (1997). We regard the assumptions (see Heckman 1997) maintained in these studies as wholly unreasonable.

Stephen L. Morgan *is a Ph.D. candidate in Sociology at Harvard University. He is currently completing his dissertation, titled "Educational Attainment and the Bayesian Dynamics of Expectation Formation." He is also working on other studies of school effects with his advisor, Aage Sørensen. With Christopher Winship, he is studying topics in quantitative methodology, and with his father, William R. Morgan, and several Nigerian colleagues, he is studying labor markets and patronage networks in northern Nigeria.*

Aage B. Sørensen *is Professor of Sociology at Harvard University and former chair of the Department of Sociology . He currently Chairs the Joint Program (Harvard Business School and the Departments of Sociology and Psychology) in Organizational Behavior. Much of his past research has focused on the study of career patterns as the movement of people among educational, occupational, and organizational slots or positions. Currently, his empirical research focuses on the impact of labor market structures on unemployment processes, on the causes and consequences of early school-leaving, and on the development of proper models for the estimation of school effects. His principal current theoretical interest is in the development of a sound theoretical basis for class analysis.*

Appendix A. The Analytic Sample

We chose an analytic sample from the NELS data set that could address the research questions in which we were interested without overwhelming those questions with all of the complexity that the NELS data offer. In this appendix, we describe the construction of the analytic sample, the further selection of a smaller subsample on which the models in Tables 3 through 5 were estimated, and the modeling of missing-data patterns.

The base year eighth-grade questionnaire was completed by 24,595 NELS respondents in 1988. To form the analytic sample, we first dropped: 6,202 base-year respondents who were randomly subsampled out before the first follow-up occurred in 1990; 2,039 base-year respondents who either dropped out of school before 1990 or were in school in 1990 but not in the tenth grade; 430 base-year respondents who were attending private high schools in 1990 that were affiliated with a religion other than Catholicism; 1,019 base-year respondents who were attending private schools in 1990 that had no religious affiliation; and 120 base-year respondents who had missing values for race. As a result, the remaining sample of base-year respondents included all potential members of our analytic sample, except for the 120 respondents who did not have valid race/ethnicity codes (nor much other valid data).[a]

The remaining 14,785 base-year respondents were classified into the nine groups for which separate "realized sample selection probabilities" were calculated (see note to Table A-1). While each NELS eighth grader can be classified as a member of only one of nine states, each eighth grader has a positive predictive probability of entering each of them. To obtain probabilities predicted from eighth-grade characteristics, we estimated a multinomial logit model with the most common destination "in same school for tenth and twelfth grade and completed all math tests" as the base category. Missing values for all independent variables were imputed with best-subset regression.

Thirty-three independent variables were specified

[a] Because school administrators could remove from the sampling frame any student they felt could not complete the base-year questionnaire and tests, there is undercoverage bias in the NELS base-year sample for eighth graders with poor English skills and learning disabilities. We made no attempt to correct for this undercoverage bias.

Appendix Table A-1. Correlations among Realized Selection Probabilities for Base Year Respondents: National Education Longitudinal Study of 1988

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	Number of Respondents
(1) SSMAT	—	—	—	—	—	—	—	—	9,400
(2) SSNMAT	–.354	—	—	—	—	—	—	—	2,641
(3) CSMAT	–.309	–.029	—	—	—	—	—	—	510
(4) CSNMAT	–.616	.215	.194	—	—	—	—	—	736
(5) NIGMAT	–.528	–.248	.196	.195	—	—	—	—	114
(6) NIGNMAT	–.660	.076	.106	.322	.514	—	—	—	102
(7) DOMAT	–.649	–.224	.204	.151	.515	.468	—	—	342
(8) DONMAT	–.810	–.108	.119	.338	.608	.576	.751	—	745
(9) STATUK	–.638	.061	.160	.441	.441	.590	.431	.463	195
Total									14,785

Notes: See Table 1 (pp. 665–66) for definitions of variable labels. Data are weighted.

for the multinomial logit: mother's education, father's education, mother's occupational prestige, father's occupational prestige, family income, reading test score, science test score, history test score, educational expectations, parents' educational expectation for student, student report of disciplinary record, parent report of student's disciplinary record, two indices of parental involvement in student's education, and dummy variables for sex, race, region, urbanicity, and family composition. Beyond the eight main-effect constants, the multinomial logit contains an additional 264 parameters. Taking twice the difference between the log-likelihoods of the unconstrained and the constants-only models yields a test statistic of 2,659—a value that is far into the tail of a chi-squared distribution with 264 degrees of freedom. Thus, while the model has no claim to parsimony, as many of the confidence intervals for the slope coefficients include zero, from the perspective of overall model fit (and our interest in obtaining selection probabilities with large variance), the likelihood-ratio test indicates that the 264 additional parameters are justifiable.[b]

We do not report all of the coefficients for the full model. In Table A-1, however, we present a correlation matrix of predicted probabilities for respondents of entering each of the nine states. Each probability is a separate nonlinear function of the same independent variables and can be interpreted as a propensity score (Rosenbaum and Rubin 1983). However, as described below, we do not use the propensity score to match cases (Smith 1997) but instead as a control function (Heckman and Robb 1985) in our math-gains models.

The correlation matrix of these probabilities (or "propensities") allows for an informal assessment of the similarity of alternative destinations as predicted by the characteristics of eighth graders. The probability of remaining in the analytic sample (SSMAT: same school and completed math tests) is negatively correlated with all other probabilities to varying degrees. The negative correlation is strongest with DONMAT, dropout and did not complete math tests, at –.810. However, SSMAT is negatively correlated with all other dropout, not-in-grade, and status unknown probabilities at a level of at least –.5. Only the negative correlations with "remaining in the same school but not completing all math tests" and "changing schools and completing all math tests" are weak to moderate, at –.354 and –.309 respectively. As a result, and after some experimentation, we concluded that summarizing selection into these nine categories with only the SSMAT probability is satisfactory.

The portion of the analytic sample used to estimate the math-gains models presented in Tables 3 through 5 includes only those respondents who remained in the same school and who completed math tests in all years. Of the 9,400 respondents in 1,000 schools, a further 159 respondents in 102 schools were excluded from the final analysis because missing student-level data prevented the construction of aggregated school-level parental involvement and/or social closure variables.

Even though the 9,241 respondents in 898 schools all ended up in the category for which SSMAT is a prediction, they were not equally likely to have done so, as predicted by their eighth-grade characteristics. We therefore included a nonlinear parameterization of SSMAT as a set of right-hand side variables in the math-gains models in order to model the missing data mechanism. When used as a control function, the nonlinear function of the propensity score orthogonalizes that portion of the error term that is otherwise correlated with the covariates because of the missing data mechanism, assuming that the missing data mechanism has been properly modeled and specified.

[b] See Rubin and Thomas (1996) for an argument against "trimming" models that estimate propensity scores.

In all models reported in Tables 3 through 5, SSMAT is coded as an orthogonal polynomial of degree three. In other words, the probability is coded with a flexible nonlinear parameterization–one linear term, one quadratic term, and one cubic term. As a result, no restrictive linearity assumption is made about the net association between math gains and the propensity to have been included in the complete data subsample. In all models, the probability of being in the final sample is positively related to math gains, indicating that students who remained eligible for the analysis were qualitatively different from those who did not.

How would the *exclusion* of SSMAT or one of its higher order parameters affect the reported coefficients in Tables 3 through 5? In general, the negative effect of tenth-grade math score on math gains would be slightly less negative. The parents' education and parents' occupation covariates and the family income covariate, at both the student and school levels, would have larger positive coefficients. Other coefficients would be nearly unaltered (in comparison with their standard errors). For comparison with our preferred Model 4 of Table 3, the coefficients for *Catholic school*, *parents have adequate say*, *friends in school*, and *parents know parents*, respectively, are: (1) 1.730, .536, .378, and –.312 without the cubic term for SSMAT; (2) 1.674, .549, .382, and –.301 with SSMAT specified as a linear probability term alone; and (3) 1.506, .513, .392, –.246 without any terms for SSMAT. All three sets of alternative coefficient estimates are within one standard error of the estimates reported for Model 4 in Table 3.

For a comparison of Model 4 from Table 3 to a supplementary (unreported) model without the propensity score, the three parameters for SSMAT are associated with a likelihood-ratio statistic of 136.6. This value is well beyond any reasonable critical value chosen from a chi-squared distribution with three degrees of freedom. Moreover, the inclusion of the three terms also decreases the true student-level variance in mathematics gains by 1.5 percent at a minimum because a substantial portion of the estimated student-level variance must be measurement error. For these two reasons, the three polynomial terms for SSMAT were retained for all models.

Appendix B: The Choice of Model

Equation 1 (see page 668) is a simple multilevel extension of the lagged models of learning used by Coleman and his colleagues (see Coleman and Hoffer 1987; Hoffer, Greeley, and Coleman 1985). There are several ways of generating this formulation, and we discuss three of them here: the traditional regressor variable method, a linear differential equation derivation, and an autoregressive distributed lag model from econometrics.

In education research, the lagged model for the change in a variable from time t_1 to time t_2 with the value of the variable at time t_1 as a right-hand side variable is referred to as the *regressor variable* method of studying change (Cronbach and Furby 1970). It was long the preferred method for analyzing change because it dealt with the problem of regression toward the mean that appears especially serious in the presence of the well-known unreliability of change scores. The main alternative, long considered inferior, is to omit the time t_1 variable from the right-hand side of equation 1 so that the amount of change is directly caused by a set of covariates and is independent of the level of the variable under consideration. Known as the *change score* method, the model can be written as:

$$Math12_{ij} - Math10_{ij} = d + cC_j + \mathbf{a'X}_j + \mathbf{b'X}_{ij} + u_j + e_{ij}, \qquad \text{(B-1)}$$

by removing $Math10_{ij}$ as an independent variable from equation 1. For our purposes, the main substantive difference between these two methods of measuring learning is that the regressor variable method assumes that learning between the tenth and the twelfth grade depends on how much a student knows when achievement is measured in the tenth grade. In contrast, the change score method assumes that the gain is independent of how much a student knows in the tenth grade. Consensus over the superiority of equation 1 has been criticized by, among others, Allison (1990), who argues that the regressor variable method is inferior to the change score method when there is substantial random measurement error in learning and when there are preexisting achievement differences between treatment groups (see Judd and Kenny 1981, chap. 6; Willett 1988).

Another interpretation of equation 1 is as the solution to a differential equation model for the learning process initially proposed by Sørensen and Hallinan (1977) and further developed for the analysis of school-sector effects in Sørensen (1996). In this "opportunities for learning" model, the amount learned between two points in time is determined by student ability and effort while the rate of gain is constrained by the amount taught. A linear differential equation that expresses these simple ideas suggests an interpretation for the coefficients on the X variables in equation 1 as measures of the contribution of these variables to the student ability and effort that generates learning. The model also suggests that λ in equation 1 can be interpreted as a measure of opportunities for learning. For this paper, the derivation of equation 1 from a differential equation model provides a justification for entering stable characteristics of students and schools as independent variables in a regressor variable model.[c]

[c] The formulation also suggests that school effects due to variation in schools in the amount they try to teach can be estimated by obtaining estimates of λ for different types of schools (see Sørensen 1996). We do

The third approach to the derivation of equation 1 can be found in the econometrics literature, where equation 1 is the autoregressive form of a distributed lag model estimated from a time series or pooled cross-section of time series (see Greene 1993, chap. 18). With time-series data, these models pose serious estimation problems because the lagged variable is generally correlated across time with the error term, and as a result OLS estimates are biased and inconsistent. The bias in the lag coefficient (*Math*10, in our case) can be large. These problems are most serious in the analysis of a single time series because all of the information is time dependent. But the lagged formulation still poses estimation problems for panel studies, even those with only two time points. We could implement the econometric solutions to the estimation problems for lagged dependent variable models by using NELS eighth-grade test scores as instruments for tenth-grade test scores, but such a procedure would rest upon unsupportable assumptions about the sources of correlated relative performance on standardized tests across testing occasions.

The econometric literature also suggests using difference models to overcome autocorrelation problems in panel data. Suppose we have static models at our two time points:

$$\begin{aligned} Math10_{ij} = {} & cC_j + \mathbf{a}'\mathbf{X}_j + \mathbf{b}'\mathbf{X}_{ij} \\ & + u_j + e_{ij}, \end{aligned} \quad \text{(B-2)}$$

and

$$\begin{aligned} Math12_{ij} = {} & d + cC_j + \mathbf{a}'\mathbf{X}_j + \mathbf{b}'\mathbf{X}_{ij} \\ & + u_j + e_{ij}, \end{aligned} \quad \text{(B-3)}$$

where we center the variables so that the intercept in equation B-2 is constrained to equal 0. Subtracting equation B-2 from equation B-3 yields:

$$\begin{aligned} Math12_{ij} - Math10_{ij} = {} & d + c\Delta C_j + \mathbf{a}'\Delta\mathbf{X}_j \\ & + \mathbf{b}'\Delta\mathbf{X}_{ij} + v_{ij}, \end{aligned} \quad \text{(B-4)}$$

Equation B-4 is very similar to equation B-1 except that the independent variables are now the changes in the original variables between the two time periods. Unfortunately, we do not have any dynamic variables to use as covariates—all of our main variables are static. Some NELS students do change school sector between the tenth and twelfth grade, but too few to yield reliable estimates of the Catholic-school effect on learning.

The derivation of the difference model from equations B-2 and B-3 makes clear an important conceptual problem with the change score model represented by equation B-1. Equations B-2 and B-3 assume that the learning process has reached equilibrium at both times. In other words, because the ability and effort of students have exhausted their effects when achievement is measured at either time point, only increases in ability and effort induced by a treatment can produce a gain in scores between the two grades. This equilibrium assumption is indeed the common one in economic applications of the difference model. However, this is not a reasonable conception of the learning process if, as suggested by equation 1 and its differential equation derivation, learning is a dynamic growth process. The implication of the difference equation formulation is that observed effects of static independent variables in equation B-1 must be effects of changes in unmeasured variables correlated with the static variables.

When not in equilibrium, but in the presence of sufficient data, an even more comprehensive econometric solution is to estimate difference-in-difference models, regressing a change score from two time periods on a lagged change score from two prior time periods (and perhaps constraining λ^*, as in the equation below, to equal 0). Such a model would be a simple and flexible parameterization of the dynamic growth assumption suggested by the differential equation framework. For example, if we also had test scores from the ninth and eleventh grades, we could estimate a model such as:

$$\begin{aligned} Math12_{ij} - Math11_{ij} = {} & d + c\Delta C_j \\ & + (\lambda^* - 1)(Math10_{ij} - \mathrm{Math9}_{ij}) \\ & + \mathbf{a}'\Delta\mathbf{X}_j + \mathbf{b}'\Delta\mathbf{X}_{ij} + v_{ij}. \end{aligned} \quad \text{(B-5)}$$

Unfortunately, we do not have ninth and eleventh grade math test scores. And as noted for equation B-4, we do not have any dynamic predictor variables, which would have to be differenced twice for equation B-5 even if we had them.

But since we do have eighth-grade math scores, we estimated what we call a "difference-in-difference change score model" with static covariates to satisfy our curiosity. To form a model similar to equation B-5, we substituted the difference between the tenth-grade and eighth-grade math scores for the tenth-grade math score in equation 1, yielding:

$$\begin{aligned} & Math12_{ij} - Math10_{ij} = d + \\ & \quad (\lambda^* - 1)(Math10_{ij} - \mathrm{Math8}_{ij}) + cC_j \\ & \quad + \mathbf{a}'\mathbf{X}_j + \mathbf{b}'\mathbf{X}_{ij} + u_j + e_{ij}, \end{aligned} \quad \text{(B-6)}$$

where $Math8_{ij}$ is the eighth-grade math score. There are several problems with the model specified in equation B-6. The substantive interpretation of the estimates of **a**, our primary interest, is complicated because eighth-grade instruction does not take place in the high schools to which these fixed effects apply. Thus, these models do not suggest easily interpretable "school effects." Moreover, equation B-6 is even more susceptible to treatment-effect bias from regression to the mean because the lagged change score $Math10_{ij} - Math8_{ij}$ is less reliable than $Math10_{ij}$ on its own.

Despite the statistical limitations of the learning model that is the core of equation 1, all of the alternative estimates obtained from estimating versions

not pursue this analysis here because we would like to use models that allow for direct comparison with the research of others on these matters–especially Coleman and his associates–and none of these others has pursued the differential equation interpretation of the regressor variable method.

of equation B-1 and equation B-6 were remarkably similar. Therefore, only those estimates obtained from the estimation of equation 1 are presented here. A set of alternative estimates based on separate estimation of models written as equation B-1 and equation B-6 are included in a supplementary appendix available from the authors on request.

REFERENCES

Allison, Paul D. 1990. "Change Scores as Dependent Variables in Regression Analysis." *Sociological Methodology* 20:93–114.

Baron, James N. and Michael T. Hannan. 1994. "The Impact of Economics on Contemporary Sociology." *Journal of Economic Literature* 32:1111–46.

Bryk, Anthony S., Valerie Lee, and Peter Holland. 1993. *Catholic Schools and the Common Good.* Cambridge, MA: Harvard University Press.

Bryk, Anthony S. and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models.* Newbury Park, CA: Sage.

Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition.* Cambridge, MA: Harvard University Press.

———. 1997a. "The Contingent Value of Social Capital." *Administrative Science Quarterly* 42:339–65.

———. 1997b. "A Note on Social Capital and Network Content." *Social Networks* 19:355–73.

———. 1998. "The Gender of Social Capital." *Rationality and Society* 10:5–46.

Carbonaro, William J. 1998. "A Little Help from My Friend's Parents: Intergenerational Closure and Educational Outcomes." *Sociology of Education* 71:295–313.

Chubb, John E. and Terry M. Moe. 1990. *Politics, Markets, and America's Schools.* Washington, DC: Brookings.

Coleman, James S. 1961. *The Adolescent Society.* New York: Free Press.

———. 1987. "The Relations between School and Social Structure." Pp. 177–204 in *The Social Organization of Schools: New Conceptualizations of the Learning Process*, edited by M. T. Hallinan. New York: Plenum.

———. 1988a. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94:S95–120.

———. 1988b. "'Social Capital' and Schools." *Education Digest* 53:6–9.

———. 1990. *Foundations of Social Theory.* Cambridge, MA: Harvard University Press.

———. 1995. "Achievement-Oriented School Design." Pp. 11–29 in *Restructuring Schools: Promising Practices and Policies*, edited by M. T. Hallinan. New York: Plenum.

Coleman, James S. and Thomas Hoffer. 1987. *Public and Private Schools: The Impact of Communities.* New York: Basic Books.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. *High School Achievement: Public, Catholic, and Private Schools Compared.* New York: Basic Books.

Cronbach, Lee J. and L. Furby. 1970. "How We Should Measure Change—Or Should We?" *Psychological Bulletin* 74:32–49.

Figlio, David N. and Joe A. Stone. 1997. "School Choice and Student Performance: Are Private Schools Really Better?" Discussion Paper No. 1141-97, Institute for Research on Poverty, University of Wisconsin, Madison, WI.

Furstenberg, Frank F., Jr. and Mary Elizabeth Hughes. 1995. "Social Capital and Successful Development Among At-Risk Youth." *Journal of Marriage and the Family* 57:580–92.

Goldstein, Harvey. 1995. *Multilevel Statistical Models.* 2d ed. New York: Halsted.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–80.

———. 1974. *Getting a Job: A Study of Contacts and Careers.* Chicago, IL: University of Chicago Press.

Greene, William H. 1993. *Econometric Analysis.* 2d ed. Englewood Cliffs, NJ: Prentice-Hall.

Hambleton, Ronald K., H. Swaminathan, and H. Jane Rogers. 1991. *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage.

Heckman, James J. 1997. "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources* 32:441–62.

Heckman, James J. and Richard Robb. 1985. "Alternative Methods for Evaluating the Impact of Interventions." Pp. 156–245 in *Longitudinal Analysis of Labor Market Data*, edited by J. J. Heckman and B. Singer. Cambridge, MA: Cambridge University Press.

Hoffer, Thomas, Andrew M. Greeley, and James S. Coleman. 1985. "Achievement Growth in Public and Catholic Schools." *Sociology of Education* 58:74–97.

Hoxby, Caroline M. 1996. "The Effects of Private School Vouchers on Schools and Students." Pp. 177–208 in *Holding Schools Accountable: Performance-Based Reform in Education*, edited by H. F. Ladd. Washington, DC: Brookings Institution.

Judd, Charles M. and David A. Kenny. 1981. *Estimating the Effects of Social Interventions.* New York: Cambridge University Press.

Lee, Valerie E., Julia B. Smith, and Robert G. Croninger. 1997. "How High School Organization Influences the Equitable Distribution of Learning in Mathematics and Science." *Sociology of Education* 70:128–50.

Longford, Nicholas T. 1993. *Random Coefficient Models*. Oxford, England: Clarendon.
Lord, Frederic M. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum.
McLanahan, Sara and Gary Sandefur. 1994. *Growing up with a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard University Press.
Morgan, William R. 1983. "Learning and Student Life Quality of Public and Private School Youth." *Sociology of Education* 56:187–202.
Murnane, Richard, Stephen E. Newstead, and Randall J. Olsen. 1985. "Comparing Public and Private Schools: The Puzzling Role of Selectivity Bias." *Journal of Business and Economic Statistics* 3:23–35.
Neal, Derek. 1997. "The Effects of Catholic Secondary Schooling on Educational Achievement." *Journal of Labor Economics* 14:98–123.
Podolny, Joel M. and James N. Baron. 1997. "Resources and Relationships: Social Networks and Mobility in the Workplace." *American Sociological Review* 62:673–93.
Portes, Alejandro. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 22:1–24.
———, ed. 1995. *The Economic Sociology of Immigration: Essays on Networks, Ethnicity and Entrepreneurship*. New York: Russell Sage Foundation.
Portes, Alejandro and Patricia Landolt. 1996. "The Downside of Social Capital." *The American Prospect* 26:18–21.
Portes, Alejandro and Ruben G. Rumbaut. 1996. *Immigrant America: A Portrait*. Berkeley, CA: University of California Press.
Portes, Alejandro and Julia Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology* 98:1320–50.
Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
Rock, Donald A. And Judith M. Pollack. 1995. *Psychometric Report for the NELS:88 Base Year (1988) through Second Follow-Up (1992)*. Washington, DC: National Center for Education Statistics.
Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.
Rubin, Donald B. and Neal Thomas. 1996. "Matching Using Estimated Propensity Scores: Relating Theory to Practice." *Biometrics* 52: 249–64.
Sandefur, Rebecca L. and Edward O. Laumann. 1998. "A Paradigm for Social Capital." *Rationality and Society* 10:481–501.
Smith, Herbert L. 1997. "Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies." *Sociological Methodology* 27:325–53.
Sørensen, Aage B. 1996. "Educational Opportunities and School Effects." Pp. 207–25 in *James S. Coleman*, edited by J. Clark. London, England: Falmer.
Sørensen, Aage B. and Maureen T. Hallinan. 1977. "A Reconceptualization of School Effects." *Sociology of Education* 50:522–35.
Stanton-Salazar, Ricardo D. and Sanford M. Dornbusch. 1995. "Social Capital and the Reproduction of Inequality: Information Networks among Mexican-Origin High School Students." *Sociology of Education* 68:116–35.
Tyack, David B. 1974. *The One Best System: A History of American Urban Education*. Cambridge, MA: Harvard University Press.
U.S. Dept. of Education, National Center for Education Statistics. 1996. *National Education Longitudinal Study: 1988–94* [CD-ROM]. Washington, DC: Office of Educational Research and Improvement [producer, distributor].
Willett, John B. 1988. "Questions and Answers in the Measurement of Change." *Review of Research in Education* 1988–89:345–422.
Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Knoft.
Winship, Christopher. 1998. "Multicollinearity and Model Misspecification: A Bayesian Analysis." Paper presented at the winter meeting of the American Sociological Association's Methodology Section, April 5, Chicago, IL.
Woolcock, Michael. 1998. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework." *Theory and Society* 27:151–208.